数理解析研究所講究録
179巻 1973年 194-205

学習可能性の理論

江原暉将*) 相沢輝昭*)
上坂吉則**) 尾関和彦*)
*) NHK総合技術研究所
**) NHK放送科学基礎研究所

§1. はじめに

学習の問題は、より一般的な枠組の中で、知識の習得に関する問題として哲学において論じられており、歴史的には対照的な2つの学説、すなわち経験論と合理論とを区別することができる[1]。この2つの見解の相違は、極端ないい方をすれば、つぎのようになる。イギリスの経験哲学の開拓者であるJ. Locke やD. Hume ら経験論者は感覚の受容器や分析器などのある種の末しょう的データ処理機構と連合、般化・分化などの初等的な帰納原理だけが人間に生得的に備っていて、たんに感覚印象を受動的に記録し、それらに帰納原理を適用して外界についての知識を得るに過ぎないと主張する。つまり、知識は、すべて外界の刺激パターンに接するという"経験"によって習得されると考える。一方、R. Descartes や

G. W. Leibniz らの合理論者は、精神（つまり"理性"―"合理論"という語はこれに由来している）が知識の唯一の源泉であると主張する。したがって、末しょう的データ処理機構や帰納原理に加えてさまざまの"観念"や"原理"が生得的に備わっていて、これらが作用して刺激パターンが"知識"にまで至り、外界が理解されると考える。

永い間にわたってのこの哲学上の論争は心理学にも引き継がれ、学習理論においても対照的な二つの学説，すなわち，経験論に由来する連合説と合理論に沿う認知説を区別することができる[2]。

連合論者は刺激パターンとそれに対する反応を連合させる仕組みこそ知識の習得，すなわち，学習の本質であるとして連合の成立する条件をもっぱら分析する。その結果，たとえばC. L. Hullの"強化の法則"などが得られている。ここで人工知能，とくにパターン認識や制御理論の分野に目を向けてみよう。そこでの学習理論の主流，すなわち統計的学習理論[3][4]は統計学における推定や検定の手法にその数理技術的な基礎を置いているが、学習過程のアルゴリズムの形式は基本的にはパーセプトロン[5]のそれに依っている。パーセプトロンの理論では刺激パターンと反応を"強化"によって"連合"させるような学習のアルゴリズムが主に研究される，

この意味で人工知能の分野でのこれまでの学習理論は，経験論の見解に立って学習システムを作り上げようとする工学的な試みと見ることができる。

一方，認知論者は生得的機構が刺激パターンによって活性化され，ために外界の“構造”を認知することが学習の本質であり，学習後の行動の変化はその認知内容によって規定されると主張する。構造の認知を重要視するこの傾向は近年しだいに高まり，連合説の認知説への接近も見られる[2]。こうした状況の中で，N. Chomsky らの言語習得に関する見解[6][7]が抬頭し，「言語学が人間精神の研究に重要な貢献をすることができると同時に，合理論者との哲学論争において，まさに一方の立場（前者）を支持する証拠を提供できる」[8]とされる。

言語は人間の精神の問題を論ずるのに最適な材料であるという思想[9]や，図形認識に本質的と思われる問題を提起しているゲシュタルト説が認知論の立場にあるという事実[2]，あるいは上で述べた連合説の認知説への接近や人工知能の分野での合理論的な学習理論の欠如を思い起すと，一般認識能力の習得という問題を合理論の立場で考察することは十分魅力のある試みと思われる。Chomsky が合理論を支持する証拠として上げた中で最も重要なのは，「言語のもつ最も顕著な

特徴の一つであり、言語使用と言語習得に関する心理学的理論の開発に対して，特に研究しがいのある問題を提供する」[8]ところの言語の"創造性"である。このことを一般認識能力の習得という枠組の中で解釈すれば、次のようになる。事物の認識者がかつて一度も学んではいないことを驚くほど大量に知っているという明白な事実 — これは経験論的見解から到底説明できない事柄である — が，"認識の創造性"換言すれば刺激パターンが知識に至るために存在する組識化原理（生得的観念）という概念によって解明され得るという。もしこの見解を認めるならば、それではどのような生得的観念のもとで、刺激パターンが形成する外の世界の構造を認知できるかという問題が生ずる。この意味での"学習の可能性"を論ずることがこの論文の目的である。学習可能性の理論とは、平易に言えば，"一を聞いて十を知る"ことのできる認識の不思議を明らかにしようとする試みといってよい。こうした合理論的観点から学習理論を展開することは、学習機械や認識装置を作りだそうとする工学的な試みに対して一つの新しい示唆を与えるに違いないと信ずる。

§2. 学習可能性の概念

はじめにいくつかの記法を用意する。Nを正整数の集合と

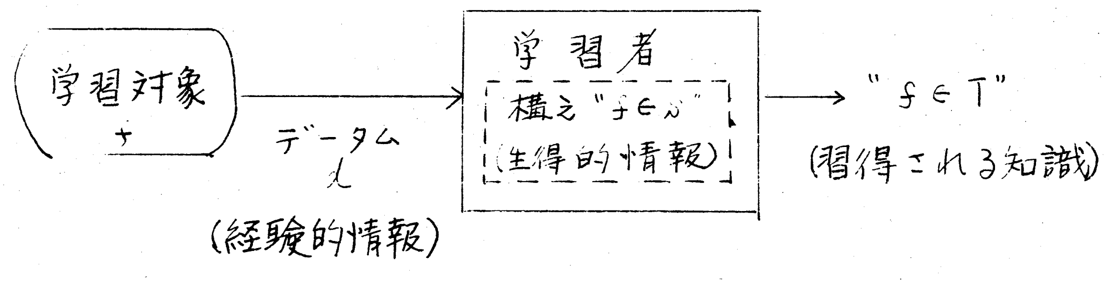
し，Fによって N から N への写像の全体を表わす。2つの N の直積集合を N^2 で表わす。

定義1．$S \subseteq N^2$ の任意の要素 (x, y)，(x', y') に対して $x = x'$ のとき $y = y'$ ならば，S は一価であるという。$f \in F$ とし，S を N^2 の一価な部分集合とする。各 $(x, y) \in S$ に対して $y = f(x)$ ならば，f は S の上にあるという。f を N^2 の部分集合とみなすとき，f が S の上にあることと f が S を含むことは同値である。したがって f が S の上にあることを $f \supseteq S$ で表わす。S が N^2 の一価な部分集合であるとき，F の要素で S の上にあるものの全体を $\mathcal{E}(S)$ で表わす：

$$\mathcal{E}(S) = \{ g ;\ g \in F,\ g \supseteq S \}.$$

定義2．N^2 の一価な有限部分集合をデータムという。すべてのデータムの族を D で表わす。空集合 ϕ はデータムであるとする。

さて学習対象と学習者からなる学習系を考える（下図）．



学習対象 f は N から N への写像であるとする。学習者は学習対象のある性質に関する知識を習得したい。このような性質

を，一般に，数学的言明，たとえば T を F のある部分集合として " $f \in T$ " などのように表わす。

ここで学習者に対して2つの仮定を置く。すなわち (1) 問題になっている学習対象が F のある部分集合 S に属しているという知識を彼は"生得的に持っているとする；(2) 学習者は学習対象の有限個の入出力関係，すなわち学習対象がその上にあるところのデータム d についての情報を学習の過程で得ることができるとする。いいかえれば，2種類の情報，つまり " $f \in S$ " という生得的情報と，" $f \in \pi(d)$ " という経験によって得られた情報に基いて，学習者は学習対象が，たとえば " $f \in T$ " といった性質を持つと結論しようとしている。このように仮定するのである。学習者がこのような試みに原理的に成功するものであるならば，知識 " $f \in T$ " を習得する可能性がその本性として，存在するといってよいであろう。任意の $f \in F$ に対して，$f \in S$ かつ $f \in \pi(d)$ ならば $f \in T$ であるということは，明らかに $S \cap \pi(d) \subseteq T$ と等価であるから，次の定義が妥当であろう。

定義3．$f \in F$ とし，S，$T \subseteq F$ とする。$f \in S \cap \pi(d) \subseteq T$ となるデータム d が存在するとき，知識 " $f \in T$ " は生得的情報 " $f \in S$ "（あるいは，単に，構え S ）のもとで学習可能であるという。

定義4. $f\in F$, $S\subseteq F$ とする。$S\cap\pi(d)=\{f\}$ となるデータム d が存在するとき，写像 f は生得的情報"$f\in S$"（あるいは単に構えS）のもとで学習可能であるという.

論を進めるに当って，ここでデータムに関するいくつかの性質を用意しておこう。証明は簡単なので省略する(10)。

補題1. d_1, d_2, $\cdots\in D$（すべてのデータムの集合）とする。このとき，

(1) $d_1\supseteq d_2\Rightarrow\pi(d_1)\subseteq\pi(d_2)$;

(2) $\pi(d_1)\cup\pi(d_2)\cup\cdots\subseteq\pi(d_1\cap d_2\cap\cdots)$;

(3) $d_1\cup d_2\cup\cdots$ が一価ならば，$\pi(d_1)\cap\cdots=\pi(d_1\cup d_2\cup\cdots)$ そうでないときは，$\pi(d_1)\cap\cdots=\phi$.

§3. 生得的情報の基本的性質

つぎの2つの定理は構えのある種の閉性を与える。証明はいずれも容易なので省略する(10).

定理1. $f\in F$; S, S_1, S_2, $T\subseteq F$ とする。

(1) 知識"$f\in T$"が S_1 のもとで学習可能であり，かつ $f\in S_2\subseteq S_1$ ならば，S_2 のもとでも学習可能である。

(2) "$f\in T$"が S_2 のもとで学習可能でなく，かつ $f\in S_2\subseteq S_1$ ならば，S_1 のもとでも学習可能でない。

(3) "$f\in T$"が S のもとで学習可能ならば，S^c のもとでは学習可能ではない。

定理2. $f\in F$; S_1, S_2, $T\subseteq F$ とする。

(1) 知識"$f\in T$"が S_1 と S_2 の双方のもとで学習可能なら

ば，$S_1 \cup S_2$のもとでも$S_1 \cap S_2$のもとでも学習可能である。

(2) $\mathcal{S}$を"$f \in T$"が学習可能であるような構えのある族とする。各$S \in \mathcal{S}$に対してあるデータムd_Sが存在して$\bigcup_{S \in \mathcal{S}} d_S$がデータムであり，かつ，すべての$S \in \mathcal{S}$に対して$f \in S \cap \pi(d_S) \subseteq T$ならば，"$f \in T$"は$\bigcup_{S \in \mathcal{S}} S$および$\bigcap_{S \in \mathcal{S}} S$のいずれのもとでも学習可能である。

§4、学習可能性の位相的側面

この節では、Fを位相づけることによって，学習可能性の概念が位相数学的用語[12]によって良く表現され得ることを示す。

$\mathcal{O}^*$をϕと$\pi(d)$達（ただし，$d \in D$）からなるFの部分集合の族とする：

$$\mathcal{O}^* = \{\phi\} \cup \{\pi(d);\ d \in D\} .$$

容易に分るように，$\mathcal{O}^*$は開集合の基底の公理を満たす；(1) 各$f \in F$に対して，$f \in W$となる$W \in \mathcal{O}^*$が存在する ；(2) 任意の$W_1, W_2 \in \mathcal{O}^*$と任意の$f \in W_1 \cap W_2$に対して；$f \in W_3 \subseteq W_1 \cap W_2$となる$W_3 \in \mathcal{O}^*$が存在する。したがって

$$\mathcal{O} = \{\ O;\ O = \bigcup_\alpha W_\alpha,\ W_\alpha \in \mathcal{O}^*\}$$

とおくと，$\mathcal{O}$は開集合系となり，Fに位相が導入される。この位相空間を学習空間と呼び，(F, $\mathcal{O}$) あるいは単にFと書く．

この位相に関して次の2つの定理が成り立つが これはこの理論で最も重要な結果である。

定理3. $f\in F$ とし, S, $T\subseteq F$ とする。知識"$f\in T$"が構えSのもとで学習可能であるのは $(F, \mathcal{O})$ に対するSの相対位相で$T\cap S$がfの近傍となるとき, かつそのときに限る。

(証明) 必要性. 仮定によって $\exists d\in D$, $f\in S\cap\pi(d)\subseteq T$。したがって $f\in S\cap\pi(d)\subseteq T\cap S$。$\pi(d)$ が $(F, \mathcal{O})$ で開であることに注意すれば $S\cap\pi(d)$ はSの相対位相で開となる。よって$T\cap S$はその相対位相でfの近傍である。

十分性. 仮定によって, $\exists O\in\mathcal{O}$, $f\in S\cap O\subseteq T\cap S$. Oは$\mathcal{O}^*$の要素によって $O=\cup_\alpha\pi(d_\alpha)$ と書けるから, $f\in\cup_\alpha(S\cap\pi(d_\alpha))\subseteq T\cap S$ となる。よって $\exists d_\beta\in D$, $f\in S\cap\pi(d_\beta)\subseteq T\cap S\subseteq T$, すなわち "$f\in T$" は$S$のもとで学習可能。

(証明終)

定理4. $f\in F$; $S\subseteq F$とする. 写像fが構えSのもとで学習可能であるのは、fがSの孤立点であるとき、かつ、そのときに限る.

(証明) 十分性. fはSの孤立点であるから, $f\in A$ かつ $S-\{f\}\subseteq\bar{A}$ な開集合Aが存在する。開集合Aは$\mathcal{O}^*$の要素によって $A=\cup_\alpha\pi(d_\alpha)$ と書け. したがって, $\exists d_\beta\in D$, $f\in\pi(d_\beta)$。

$f \in S$ に注意して，$f \in S \cap \pi(d_\beta)$。次に$S \cap \pi(d_\beta) \subseteq \{f\}$を示そう。もしそうでないとすると，$g \in S \cap \pi(d_\beta)$ かつ $g \neq f$ なる写像gが存在するはずである。$S-\{f\} \subseteq A$ であったから，$g \in S-\{f\} \subseteq A = \cap_\alpha \pi(d_\alpha)^c \subseteq \pi(d_\beta)^c$。すなわち $g \notin \pi(d_\beta)$ を得るが，これは$g \in S \cap \pi(d_\beta)$に矛盾する。よって，結局，$S \cap \pi(d_\beta) = \{f\}$，すなわち，$f$は$S$のもとで学習可能。

必要性．fがSのもとで学習可能であるから，$^\exists d \in D$，$S \cap \pi(d) = \{f\}$。よってまず$f \in S$を得る。次に，$f \notin A$ かつ $S-\{f\} \subseteq A$ なる閉集合Aが存在することを示そう。$\pi(d)$ は開であるから $A = \pi(d)^c$ とおく．実際，明らかに$f \notin A$。いま$S-\{f\} \nsubseteq A$と仮定しよう。すると$g \notin A$，すなわち，$g \in \pi(d)$ なる$g \in S-\{f\}$が存在し，したがって，$g \in S \cap \pi(d)$ かつ $g \neq f$，しかし．これは$S \cap \pi(d) = \{f\}$に矛盾する。かくして証明が完了した。

（証明終）

次に，σをN^2から$\{0,1\}$への写像とする：$x, y \in N$に対して　$\sigma(x,y)=0$（$x=y$のとき），$\sigma(x,y)=1$（$x \neq y$のとき）．δをF^2の上で定義された実数値関数：$f, g \in F$に対して　$\delta(f,g) = \sum_{x=1}^{\infty} \frac{1}{2^x} \sigma(f(x), g(x))$ とする。すると，δがF上で定義された距離関数であることが容易に確かめら

れ，したがって距離空間（F, δ）を得る。この時次の定理が証明される(11)．

定理5．距離空間（F, δ）は位相空間として，学習空間に同相である。

この定理を用いて，定理4は次のように書き改めることができる。

定理4a．写像fが構えSのもとで学習可能であるのは，ある$\varepsilon>0$が存在して，f以外のすべての$g\in S$に対して$\delta(f, g)\geq\varepsilon$となっているとき，かつその時に限る。

最後に学習空間の位相的性質を列挙すると次のようになる(11)．

第2可算公理	○	完備距離空間	○
可分	○	コンパクト	×
完全不連結	○	局所コンパクト	×

§5．文献

(1) B. Russel（市井三郎訳）：西洋哲学史 みすず書房（1959）

(2) 東洋他編：学習心理学ハンドブック，金子書房（1958）．

(3) 上坂吉則：パターン認識と学習の理論，総合図書（1971）.

(4) J.M. Mendel and K.S. Fu : Adaptive learning and pattern recognition systems, Academic Press (1970).

(5) R. Rosenblatt : Principles of neurodynamics, Spartan Books (1962).

(6) N. Chomsky : Aspects of the theory of syntax, The MIT Press (1965).

(7) ―――――: Cartesian linguistics - A chapter in the history of rationalist thought, Harper & Row (1966).

(8) J. Lyons : Chomsky, Wm. Collins & Co., Ltd (1970).

(9) N. Chomsky : Language and mind, Harcourt, Brace & World Inc. (1968).

(10) 上坂吉則・相沢輝昭・江原暉将・尾関和彦：学習可能性の理論，NHK研究報告書（1972.9）.

(11) 相沢輝昭・上坂吉則・江原暉将・尾関和彦：学習空間の位相的性質，NHK研究報告書（1972.9）.

(12) 河田敬義・三村征雄：現代数学概説II，岩波書店（1965）.